Computed tomography synthesis from magnetic resonance images in the pelvis using multiple random forests and auto-context features

Andreasen, Daniel; Morgenthaler Edmund, Jens; Zografos, Vasileios; Menze, Bjoern H.; Van Leemput, Koen

Published in:
Proceedings of SPIE

Link to article, DOI:
10.1117/12.2216924

Publication date:
2016

Document Version
Peer reviewed version

Computed Tomography synthesis from Magnetic Resonance images in the pelvis using multiple Random Forests and Auto-Context features

Daniel Andreasen[a,b], Jens M. Edmund[b], Vasileios Zografos[c], Bjoern H. Menze[c], and Koen Van Leemput[a,d]

[a]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark
[b]Radiotherapy Research Unit, Department of Oncology, Gentofte and Herlev Hospital, University of Copenhagen, Herlev, Denmark
[c]Department of Computer Science, Technische Universität München, Munich, Germany
[d]A.A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129, USA

ABSTRACT

In radiotherapy treatment planning that is only based on magnetic resonance imaging (MRI), the electron density information usually obtained from computed tomography (CT) must be derived from the MRI by synthesizing a so-called pseudo CT (pCT). This is a non-trivial task since MRI intensities are neither uniquely nor quantitatively related to electron density. Typical approaches involve either a classification or regression model requiring specialized MRI sequences to solve intensity ambiguities, or an atlas-based model necessitating multiple registrations between atlases and subject scans. In this work, we explore a machine learning approach for creating a pCT of the pelvic region from conventional MRI sequences without using atlases. We use a random forest provided with information about local texture, edges and spatial features derived from the MRI. This helps to solve intensity ambiguities. Furthermore, we use the concept of auto-context by sequentially training a number of classification forests to create and improve context features, which are finally used to train a regression forest for pCT prediction. We evaluate the pCT quality in terms of the voxel-wise error and the radiologic accuracy as measured by water-equivalent path lengths. We compare the performance of our method against two baseline pCT strategies, which either set all MRI voxels in the subject equal to the CT value of water, or in addition transfer the bone volume from the real CT. We show an improved performance compared to both baseline pCTs suggesting that our method may be useful for MRI-only radiotherapy.

Keywords: radiotherapy, magnetic resonance imaging, pseudo CT, CT synthesis, random forest, auto-context

1. INTRODUCTION

In recent years, the interest in eliminating the planning computed tomography (CT) from the workflow in external beam radiotherapy (RT) has increased.[1–4] Instead, the entire workflow would be based on magnetic resonance imaging (MRI), thus eliminating systematic registration uncertainties between the MRI and CT[5,6] and simplifying the treatment chain. It is a difficult task to exclude the CT, since MRI does not contain information on electron density which is needed for dose calculations. Furthermore, the signal from compact bone in conventional MRI is weak or missing due to the majority of short T_2 components.[7] This prohibits 2D-based patient setup using bone anatomy. By synthesizing a so-called pseudo CT (pCT) from the MRI scan, the above issues would be solved, making an MRI-only workflow feasible.

Several approaches for CT synthesis have been proposed and in general, the methods can be split into two categories: 1) voxel-based and 2) atlas-based. With the former, a classification or regression model is trained

Correspondence to Daniel Andreasen, E-mail: dana@dtu.dk

based on individual voxel intensities of pre-acquired MRI/CT pairs from a number of patients.[2,8–10] Other variants in this category include unsupervised methods based on MRI physics knowledge.[11,12] The voxel-based methods generally require specialized ultra short echo time MRI sequences for acquiring the bone signal, or manual delineations to overcome the bone/air intensity-ambiguity. This either prolongs the patient scan time or puts additional workload on the personnel. In recent publications, many variants of methods from the atlas-based category have been proposed.[13–19] Here, pre-acquired and co-registered MRI/CT pairs are used as atlases that describe the relation between MRI and CT. To predict a pCT, the atlas MRIs are first deformably registered with the patient MRI. The estimated deformations are then applied to the atlas CTs and finally the deformed CTs are fused to provide the pCT estimate. These methods rely on conventional MRI sequences and the accuracy is generally high. One difficulty, however, is that the use of deformable registrations can be slow, especially with multiple atlases, and may fail if the subjects are very dissimilar. Using pattern recognition based on patch similarities can remove some of the dependency on an accurate deformable registration[13] and a faster affine registration can even be used instead.[20] Still, the registration is a computationally demanding step, which must be performed for each atlas every time a pCT is predicted. Especially since the trend seems to go towards adding more atlases to the database, i.e., 15-38 patients[16,17,19] this can become a time-consuming task.

Methods using a random forest[21] (RF) have previously been proposed for predicting one MRI contrast from another.[22,23] This task differs somewhat from CT synthesis because the bone/air ambiguity does not need to be solved; the prediction model only needs to identify and predict the lack of signal in these regions, but is not required to distinguish whether it is caused by bone or air. In this work, we use RF models to predict pCTs based on conventional T_1-weighted MRI scans without registrations during pCT synthesis. We overcome intensity-ambiguities by using local texture descriptors, spatial features and edge information. Furthermore, inspired by recent works in tumor and organ segmentation,[24,25] we use the concept of auto-context (AC)[26] to iteratively learn and improve context features. We use a combination of classification RFs for learning context features and a regression RF to ultimately predict a pCT.

By far most work in CT synthesis has been on generating pCTs of the brain, although a few publications have focused on the pelvis.[8,14,16,17] This region is theoretically more challenging due to the greater inter-patient anatomical variability and because of between-scan non-rigid organ movement. We apply our method in this challenging region, and evaluate and compare the resulting pCTs to baseline pCT approaches both in terms of the voxel-wise prediction accuracy and the radiologic error.

2. METHODS

2.1 Data Collection and Pre-processing

MRI and CT scans of 10 prostate RT patients were retrospectively obtained from a Philips 1 T open MR scanner using a body coil and from a Philips Brilliance Big Bore CT scanner. T_1-weighted scans were acquired with a gradient echo sequence (TE/TR = 10/623 ms) and a voxel resolution of $0.8 \times 0.8 \times 4.4$ mm for an in-plane matrix between 528×528 and 640×640 voxels and $16 - 24$ slices. The CT scans were acquired using a standard protocol for pelvic scans (120 kVp, 232-503 mAs) with a voxel resolution between $0.78 \times 0.78 \times 2$ mm and $1.4 \times 1.4 \times 2$ mm, for an in-plane matrix of 512×512 voxels and 129-199 slices. The patients were fixed in treatment position during both the MRI and CT scanning. The T_1-weighted scans were bias field corrected using the N4 algorithm[27] after which the intensities were normalized to a common range using a histogram matching approach.[28] To generate a ground truth relationship between the MRI and CT scans, which accounts for body outline and organ deformations between the scans, the MRI scans were deformably registered and resampled to their corresponding CT using Elastix.[29] For this purpose, a multi-resolution B-spline transformation model and normalized mutual information was used. Note that this registration is carried out before training the RFs and is not a part of the pCT prediction step. Finally, in order to remove noise without severely blurring edges, a simple 2D median filter was applied slice-wise to the MRI scans.

2.2 Initial Feature Extraction

Using the MRI intensity information directly to train a prediction model may limit its general applicability because the absolute MRI intensities are irreproducible and varies between patients, scans and scanners. For this reason, we use features that to some degree are independent of linear gray-scale changes. We use Local

Binary Pattern (LBP) features[30] to capture textural information. For each voxel, we do a relational comparison of the mean value of a cuboidal region centered on that voxel with mean values of 26 cuboidal regions positioned on a sphere around that voxel. This results in a 26 digit binary number which is converted to a decimal texture feature (dubbed LBP_μ). We also incorporate a variant of the LBP where the standard deviation of the regions is compared instead (dubbed LBP_σ). We use three LBP_μ and two LBP_σ features with differing region size (RS), 3^3 - 5^3 voxels, and sphere radius (SR), 5 - 13 voxels, to capture information at different scales. As a simple edge detector, we use the standard deviation of a $3 \times 3 \times 3$ neighborhood around each voxel. As spatial features, we use the (x, y) position of each voxel relative to the center of mass of the patient and the body outline. Figure 1 shows examples of the initial features.

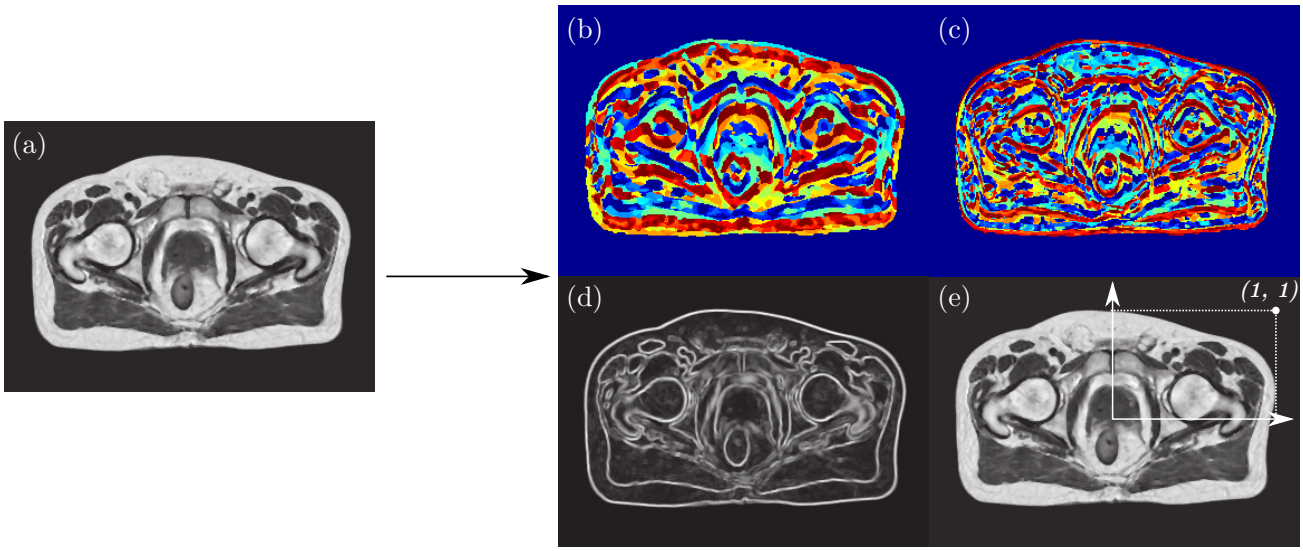

Figure 1. Features are extracted from the T_1-weighted MRI shown in (a). (b) and (c) shows examples of the LBP_μ (SR/RS = $11/5^3$ voxels) and LBP_σ (SR/RS = $7/5^3$ voxels), respectively. (d) shows the standard deviation feature. (e) shows the coordinate system used to derive the spatial features; the origin is positioned in the center of mass of the patient and the axes are scaled relative to the outer contour of the patient.

2.3 Initial Classification RF Training

In the initial step of our approach, we train a classification RF to segment an MRI scan into air, fat, soft tissue and bone classes based on the features described above. We carry out all model training in a leave-one-out manner using nine patients as training data and applying the model on the tenth. The training target segmentations are generated by thresholding the CT scans using the following criteria: air = $[-1000, -200]$ HU, fat = $]-200, 0[$ HU, soft tissue = $[0, 150]$, bone = $]150, \infty[$ HU. For each training patient, we then sample our features at 6×10^5 and 10^5 random positions within the body outline and within the bone volume, respectively. This defines our training set for the classification RF. We use an RF with 500 trees and a minimum terminal leaf size of 50 voxels. Because the bone class is relatively small compared to the others, we use a stratified sampling scheme, meaning that an equal amount of cases from each class were chosen at random during training of each tree. This is achieved by sub-sampling the bigger classes. The parameter settings were chosen as they resulted in a good compromise between the out-of-bag error rate and training time.

2.4 Auto-context Features

In the next step of our approach, we use the initial RF to create a variant of auto-context (AC) features[26] for all training patients, which are then used in combination with the initial features to train an improved classification RF. This process is repeated three times, each time using the most recently trained RF to improve the AC features. The AC features are calculated based on the three probability maps of a voxel belonging to the fat, soft tissue or bone classes. For each voxel, the feature extracts the mean probability of belonging to each class in a small region centered on that voxel, along with the same mean probabilities of six regions chosen at a fixed

distance from it in the posterior/anterior, superior/inferior and lateral directions. The AC feature extraction is illustrated in Figure 2. In this way, we encode the approximate locations and relations of the tissue classes. The AC RFs consist of 150 trees and otherwise the same parameter settings as the initial forest. For each RF at this stage, we sample the training data randomly in the same fashion as the first RF. This means that each forest is not trained on exactly the same training data as in the iteration before. Note however, that we sample with replacement so overlap can occur.

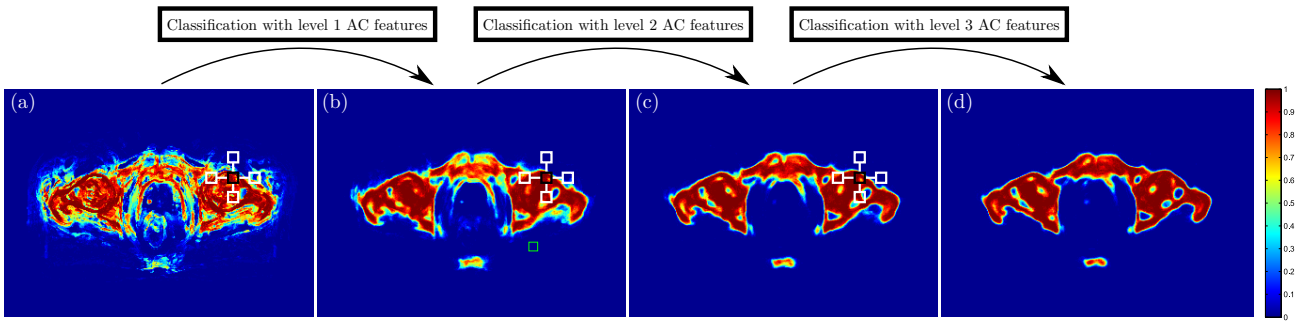

Figure 2. Examples of the output probability maps of the bone class for a test patient after sending the data through (a) the first classification RF and (b)-(d) the subsequent classification RFs with improving AC features. The AC feature for a voxel consists of the mean probability of a small region (shown in black) around the voxel along with the mean probability of 6 small regions at a fixed distance in the posterior/anterior, superior/inferior and lateral directions (shown in white for the posterior/anterior and lateral directions).

2.5 Regression Forest Training

As a final step, we train a regression RF based on the initial features and the last level of AC features. We use the non-thresholded CT scans as target values and sample our features from each training patient at 1.3×10^6 and 10^5 random positions within the body outline and bone volume, respectively. The regression forest consists of 30 trees and a minimum terminal leaf size of 70 voxels.

2.6 Benchmark Methods

As simple benchmark methods, we generate pCTs by assigning all voxels within the body outline of the MRI a bulk density value of 0 Hounsfield units (HU) corresponding to water ($\mathrm{MRI_w}$). Furthermore, we generate a pCT which in addition contains the bone volume transferred from the real CT and assigned a bulk density of the mean HU in that volume ($\mathrm{MRI_{wb}}$). This mimics a method where the bone has been manually (and perfectly) delineated on the MRI.

2.7 Evaluation

Similar to our previous work,[20] we calculate the voxel-wise mean absolute error ($\mathrm{MAE_{vox}}$) defined as:

$$\mathrm{MAE_{vox}} = \frac{1}{N}\sum_{i=1}^{N} |\mathrm{CT}_i - \mathrm{pCT}_i|, \tag{1}$$

where N is the number of voxels inside the body outline, CT(i) and pCT(i) are the HU value at voxel i inside the body outline of the real CT or pCT, respectively. Similarly, we calculate the mean error, $\mathrm{ME_{vox}}$. To provide a radiologic error measure that takes into account the attenuation properties of the tissue,[31] we also calculate the mean absolute error in water-equivalent path lengths ($\mathrm{MAE_{WEPL}}$). This is calculated along radial spokes from the center of mass of the patient to the body surface. The WEPL, l', of a spoke is calculated as:

$$l' = \sum_i \Delta l_i \times \rho_i, \tag{2}$$

with Δl_i being the physical path length traveled by a photon in voxel i and ρ_i its electron density relative to water. The value of ρ_i is found from the voxel HU value using a verified standard lookup table in the treatment planning system (Eclipse v11.0, Varian Medical Systems Inc., Palo Alto, CA). We calculate l' in 800 radial spokes towards the body outline in both the real CT and pCT. The $\mathrm{MAE}_{\mathrm{WEPL}}$ is then defined as:

$$\mathrm{MAE}_{\mathrm{WEPL}} = \frac{1}{L}\sum_{j=1}^{L} |l_j'^{CT} - l_j'^{pCT}|, \tag{3}$$

where L is the number of spokes and l'_j is the WEPL of the jth spoke in either the CT or the pCT. Similarly we calculate the mean error in WEPL, $\mathrm{ME}_{\mathrm{WEPL}}$. We perform paired t-tests to test for significant differences in the error metrics between the proposed and the baseline methods. $p < 0.05$ is considered significant.

3. RESULTS

In Figure 3, a visual comparison of the different pCTs and the real CT is shown for a representative transverse slice of one patient. The proposed method is able to capture the challenging bone region but parts of the dense bone is still underestimated as seen in the difference map (Figure 3(f)).

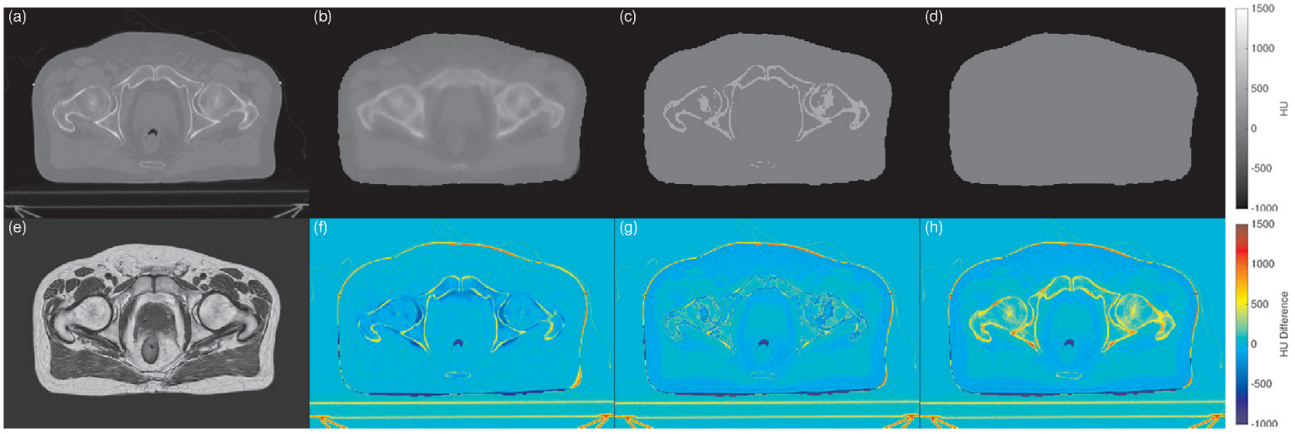

Figure 3. Transverse slices of the real CT and pCTs for one patient. (a) and (e) show the real CT and corresponding T_1-weighted MRI. (b)-(d) show the pCT generated with our method, the $\mathrm{MRI}_{\mathrm{wb}}$ and the $\mathrm{MRI}_{\mathrm{w}}$, respectively. (f)-(g) show the difference in HU between the real CT and the pCTs in (b), (c) and (d), respectively. No colorbar is shown for the MRI image.

Table 1 shows the results of the voxel-wise and radiologic evaluation. The proposed method performs well in comparison to the baseline methods with significantly lower voxel-wise and radiologic absolute errors. The $\mathrm{ME}_{\mathrm{vox}}$ reveals an overall tendency of the proposed method to overestimate the HU value. There is also an average overestimation of the WEPL, revealed by the negative $\mathrm{ME}_{\mathrm{WEPL}}$.

	Proposed	$\mathrm{MRI}_{\mathrm{wb}}$	p-value	$\mathrm{MRI}_{\mathrm{w}}$	p-value
MAE_{vox} [HU]	58 ($\sigma = 9$)	90 ($\sigma = 11$)	$< 10^{-6}$	105 ($\sigma = 12$)	$< 10^{-7}$
ME_{vox} [HU]	-10 ($\sigma = 10$)	-41 ($\sigma = 12$)	$< 10^{-5}$	-16 ($\sigma = 11$)	0.06
$\mathrm{MAE}_{\mathrm{WEPL}}$ [mm]	1.3 ($\sigma = 0.4$)	2.0 ($\sigma = 0.3$)	$< 10^{-3}$	2.5 ($\sigma = 0.3$)	$< 10^{-6}$
$\mathrm{ME}_{\mathrm{WEPL}}$ [mm]	-0.7 ($\sigma = 0.6$)	-0.4 ($\sigma = 0.5$)	0.05	0.6 ($\sigma = 0.6$)	$< 10^{-4}$

Table 1. The voxel-wise $\mathrm{MAE}_{\mathrm{vox}}$ and $\mathrm{ME}_{\mathrm{vox}}$ in HU and the $\mathrm{MAE}_{\mathrm{WEPL}}$ and $\mathrm{ME}_{\mathrm{WEPL}}$ of the water equivalent path lengths (in mm). Average value and standard deviation (σ) for the 10 patients are shown for the different methods along with the p-value from a paired t-test on the difference between the proposed method and each baseline method.

4. DISCUSSION

In this work, we explored a machine learning approach for creating a pCT of the pelvic region based on T_1-weighted MRI. The method required no specialized MRI sequences or atlases to overcome intensity ambiguities in bone/air regions of the MRI. Instead, we extracted simple textural, edge and spatial features from the MRI and trained an RF for tissue classification. As shown in Figure 2(a), the initial classification provides a rough estimate of the positions of each tissue class but some uncertainty and wrong predictions remain. The iterative refinement provided by the AC features in the next levels of RFs converges the classification towards a more correct solution. In Figure 3(b), the visual quality of the pCT comes close to the real CT. There seems to be a blurring effect in the pCT, which might be caused by the LBP and AC features only capturing information at scales larger than the voxels, i.e., from small regions. The blurring causes the bone region to smear into the soft tissue region, which might explain the average overestimation of the HU. Incorporating more features at the voxel scale could potentially remove some of the blurring effect.

The MRI_{wb} method should have a high accuracy in the bone region due to its exact geometry. However, assigning the mean HU value to the whole region means that parts of it are overestimated. On average, 65% of the voxels within the bone region are below the mean bone HU value, which illustrates the difficulties in assigning a single bulk density to segmented regions. Since, in addition, the fat voxels are not accounted for, the ME_{vox} shows an overall tendency for overestimating the HU value.

The proposed method had an average MAE_{vox} of 58 HU, which was significantly lower than the baseline methods. Other recently published methods in the multi-atlas category[16,17] reported average MAE_{vox} and ME_{vox} in the pelvic region of 36.5−40.5 and 0.6−1.9 HU, respectively. They further showed a high dosimetric accuracy of their pCTs. These methods use MRI/CT atlases from up to 38 patients and deformable registration to create pCTs. In contrast, our approach requires no atlases and no registrations at run-time, yielding a less computationally demanding method that is easier to distribute.

We have previously seen that comparing pCT methods based only on simple voxel-wise and radiologic error metrics may not be sufficient.[20] Instead an additional dosimetric evaluation should be carried out, which can highlight if the prediction fails in dosimetrically important regions or in less relevant regions. A dosimetric evaluation remains part of our future work.

Parallel to our work, another group has published a method for CT synthesis with many similarities to ours.[32] Our method differs slightly in the features used and the way auto-context is implemented. Furthermore, their method is still dependent on rigid registrations during pCT prediction to bring their training subjects in alignment with the test patient. They use a so-called structured RF to predict full CT patches instead of just voxels and report an average MAE_{vox} of 48.1 HU for the pelvis of 22 subjects. This is better than our method, but whether it is due to a more advanced RF model or simply because of more training data is hard to say.

There are still multiple potential ways of improving the proposed method. A range of different features could be tested for improving the classification. Furthermore, the LBP and AC features used here were pre-defined in terms of the region sizes and distances between the regions. An often used strategy would be to let the RF randomly pick and learn these parameters. For efficiency, a supervoxel-based strategy could also be employed.[25]

In the current work, we used features that could make the method broadly applicable on data from different scanners. Still, the features that depend on e.g., the standard deviation may describe a specific noise level in the training scans, which may be different in other scanners. Furthermore, with the rather limited amount of data available to train our models, there is a risk of overfitting to the data from our scanner, especially as more levels of auto-context are used. Testing the method's robustness using data from various scanners as well as training models based on more data remains part of our future work.

5. CONCLUSION

We presented a method for creating synthetic CT scans from MRI using random forests and textural, spatial and edge information. Furthermore, we iteratively learned auto-context features to improve the predictive performance of our models. The method worked on conventional T_1-weighted MRI scans and required no registrations at run-time. Our results indicate an improved performance compared to simple bulk density assigned pCTs but further evaluations and comparisons with more advanced methods are still needed.

ACKNOWLEDGMENTS

This work was supported by a research grant from Varian Medical Systems, Inc.

REFERENCES

[1] Kristensen, B. H., Laursen, F. J., Løgager, V., Geertsen, P. F., and Krarup-Hansen, A., "Dosimetric and geometric evaluation of an open low-field magnetic resonance simulator for radiotherapy treatment planning of brain tumours.," *Radiotherapy and Oncology* **87**, 100–9 (4 2008).

[2] Edmund, J. M., Kjer, H. M., Van Leemput, K., Hansen, R. H., Andersen, J. A., and Andreasen, D., "A voxel-based investigation for mri-only radiotherapy of the brain using ultra short echo times.," *Physics in medicine and biology* **59**, 7501–19 (12 2014).

[3] Jonsson, J. H., Karlsson, M. G., Karlsson, M., and Nyholm, T., "Treatment planning using mri data: an analysis of the dose calculation accuracy for different treatment regions," *Radiation Oncology* **5**, 62 (1 2010).

[4] Lambert, J., Greer, P. B., Menk, F., Patterson, J., Parker, J., Dahl, K., Gupta, S., Capp, A., Wratten, C., Tang, C., Kumar, M., Dowling, J., Hauville, S., Hughes, C., Fisher, K., Lau, P., Denham, J. W., and Salvado, O., "Mri-guided prostate radiation therapy planning: Investigation of dosimetric accuracy of mri-based dose planning.," *Radiotherapy and Oncology* **98**, 330–4 (3 2011).

[5] Ulin, K., Urie, M. M., and Cherlow, J. M., "Results of a multi-institutional benchmark test for cranial CT/MR image registration," *International Journal of Radiation Oncology*Biology*Physics* **77**(5), 1584 – 1589 (2010).

[6] Nyholm, T., Nyberg, M., Karlsson, M. G., and Karlsson, M., "Systematisation of spatial uncertainties for comparison between a mr and a ct-based radiotherapy workflow for prostate treatments.," *Radiation Oncology* **4**, 54 (1 2009).

[7] Robson, M. and Gatehouse, P., "Magnetic resonance: An introduction to ultrashort te (ute) imaging," *Journal of computer assisted tomography* **27**(6), 825–846 (2003).

[8] Korhonen, J., Kapanen, M., Keyriläinen, J., Seppälä, T., and Tenhunen, M., "A dual model hu conversion from mri intensity values within and outside of bone segment for mri-based radiotherapy treatment planning of prostate cancer," *Medical Physics* **41**(1), – (2014).

[9] Johansson, A., Karlsson, M., and Nyholm, T., "CT Substitute Derived from MR Sequences with Ultrashort Echo Time," *Medical Physics* **38**(5), 2708–2714 (2011).

[10] Rank, C. M., Hünemohr, N., Nagel, A. M., Röthke, M. C., Jäkel, O., and Greilich, S., "MRI-based simulation of treatment plans for ion radiotherapy in the brain region," *Radiotherapy and Oncology* **109**(3), 414 – 418 (2013).

[11] Keereman, V., Fierens, Y., Broux, T., De Deene, Y., Lonneux, M., and Vandenberghe, S., "Mri-based attenuation correction for pet/mri using ultrashort echo time sequences.," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* **51**, 812–8 (5 2010).

[12] Berker, Y., Franke, J., Salomon, A., Palmowski, M., Donker, H. C. W., Temur, Y., Mottaghy, F. M., Kuhl, C., Izquierdo-Garcia, D., Fayad, Z. a., Kiessling, F., and Schulz, V., "Mri-based attenuation correction for hybrid pet/mri systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/dixon mri sequence.," *Journal of nuclear medicine* **53**, 796–804 (5 2012).

[13] Hofmann, M., Steinke, F., Scheel, V., Charpiat, G., Farquhar, J., Aschoff, P., Brady, M., Schölkopf, B., and Pichler, B. J., "MRI-Based Attenuation Correction for PET/MRI: A Novel Approach Combining Pattern Recognition and Atlas Registration," *Journal of Nuclear Medicine* **49**, 1875–1883 (2008).

[14] Dowling, J. A., Lambert, J., Parker, J., Salvado, O., Fripp, J., Capp, A., Wratten, C., Denham, J. W., and Greer, P. B., "An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy," *International journal of radiation oncology, biology, physics* **83**, e5–e11 (05 2012).

[15] Uh, J., Merchant, T. E., Li, Y., Li, X., and Hua, C., "MRI-based treatment planning with pseudo CT generated through atlas registration," *Medical Physics* **41**(5), – (2014).

[16] Siversson, C., Nordström, F., Nilsson, T., Nyholm, T., Jonsson, J., Gunnlaugsson, A., and Olsson, L. E., "Technical note: Mri only prostate radiotherapy planning using the statistical decomposition algorithm," *Medical Physics* **42**(10), 6090–6097 (2015).

[17] Dowling, J. a., Sun, J., Pichler, P., Rivest-Hénault, D., Ghose, S., Richardson, H., Wratten, C., Martin, J., Arm, J., Best, L., Chandra, S. S., Fripp, J., Menk, F. W., and Greer, P. B., "Automatic substitute ct generation and contouring for mri-alone external beam radiation therapy from standard mri sequences," *International Journal of Radiation Oncology*Biology*Physics* (2015).

[18] Burgos, N., Cardoso, M. J., Modat, M., Pedemonte, S., Dickson, J., Barnes, A., Duncan, J. S., Atkinson, D., Arridge, S. R., Hutton, B. F., and Ourselin, S., "Attenuation correction synthesis for hybrid pet-mr scanners," in [*Medical image computing and computer-assisted intervention: MICCAI 2013*], **16**, 147–154 (1 2013).

[19] Burgos, N., Cardoso, M. J., Guerreiro, F., Veiga, C., Modat, M., McClelland, J., Knopf, A.-c., Punwani, S., Atkinson, D., Arridge, S. R., Hutton, B. F., and Ourselin, S., "Robust ct synthesis for radiotherapy planning: Application to the head and neck region," in [*Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*], Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., eds., *Lecture Notes in Computer Science* **9350**, 476–484, Springer International Publishing (2015).

[20] Andreasen, D., Van Leemput, K., Hansen, R. H., Andersen, J. A., and Edmund, J. M., "Patch-based generation of a pseudo ct from conventional mri sequences for mri-only radiotherapy of the brain," *Medical physics* **42**(4), 1596–1605 (2015).

[21] Breiman, L., "Random forests," *Machine learning* **45**(1), 5–32 (2001).

[22] Chen, M., Jog, A., Carass, A., and Prince, J. L., "Using image synthesis for multi-channel registration of different image modalities," in [*Proc SPIE Int Soc Opt Eng.*], Ourselin, S. and Styner, M. A., eds., **136**, 94131Q (3 2015).

[23] Jog, A., Carass, A., Pham, D. L., and Prince, J. L., "Random forest flair reconstruction from t1, t2, and pd-weighted mri," in [*2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*], 1079–1082, IEEE (4 2014).

[24] Qian, C., Wang, L., Yousuf, A., Oto, A., and Shen, D., "In vivo mri based prostate cancer identification with random forests and auto-context model," in [*Machine Learning in Medical Imaging*], Wu, G., Zhang, D., and Zhou, L., eds., *Lecture Notes in Computer Science* **8679**, 314–322, Springer International Publishing (2014).

[25] Zografos, V., Valentinitsch, A., Rempfler, M., Tombari, F., and Menze, B., "Hierarchical multi-organ segmentation without registration in 3d abdominal ct images," *MCV workshop (MICCAI 2015)* (2015).

[26] Tu, Z., "Auto-context and its application to high-level vision tasks," in [*Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*], 1–8, IEEE (2008).

[27] Tustison, N. J., Avants, B. B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J. C., et al., "N4itk: improved n3 bias correction," *Medical Imaging, IEEE Transactions on* **29**(6), 1310–1320 (2010).

[28] Nyul, L., Udupa, J., and Zhang, X., "New variants of a method of MRI scale standardization," *Medical Imaging, IEEE Transactions on* **19**, 143–150 (Feb 2000).

[29] Stefan Klein and Marius Staring and Keelin Murphy and Max A. Viergever and Josien P.W. Pluim, "elastix: a toolbox for intensity-based medical image registration," *IEEE Transactions on Medical Imaging* **29**, 196 – 205 (January 2010).

[30] Ojala, T., Pietikäinen, M., and Mäenpää, T., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **24**(7), 971–987 (2002).

[31] Rank, C. M., Hünemohr, N., Nagel, A. M., Röthke, M. C., Jäkel, O., and Greilich, S., "Mri-based simulation of treatment plans for ion radiotherapy in the brain region.," *Radiotherapy and Oncology* **109**, 414–8 (12 2013).

[32] Huynh, T., Gao, Y., Kang, J., Wang, L., Zhang, P., Lian, J., and Shen, D., "Estimating ct image from mri data using structured random forest and auto-context model," *IEEE Transactions on Medical Imaging* **0062**(c), 1–1 (2015).